Bootstrap for testing the equality of selfsimilarity exponents across multivariate time series

Herwig Wendt, Charles-Gérard Lucas, Patrice Abry, Gustavo Didier

▶ To cite this version:

HAL Id: hal-03381950

https://hal.science/hal-03381950

Submitted on 18 Oct 2021

Bootstrap for testing the equality of selfsimilarity exponents across multivariate time series

Charles-Gérard Lucas, Patrice Abry
Univ Lyon, Ens de Lyon, Univ Claude Bernard,
CNRS, Laboratoire de Physique, Lyon, France.
firstname.lastname@ens-lyon.fr

Herwig Wendt
IRIT, Univ. Toulouse,
CNRS, Toulouse, France.
herwig.wendt@irit.fr

Gustavo Didier
Math. Dept., Tulane University,
New Orleans, USA.
gdidier@tulane.edu

Abstract—Because of the ever-increasing collections of multivariate data, multivariate selfsimilarity has become a widely used model for scale-free dynamics, with successful applications in numerous different fields. Multivariate selfsimilarity exponent estimation has therefore received considerable attention, with notably an original procedure recently proposed and based on the eigenvalues of the covariance random matrices of the wavelet coefficients at fixed scales. Expanding on preliminary work aiming to test for the equality of the selfsimilarity exponents in bivariate time series, we propose and study here a truly multivariate procedure that permits, from a single observation of multivariate time series, to test for the equality of several, possibly many, selfsimilarity exponents. It is based on an original bootstrap procedure, applied in a multivariate time-scale domain and designed to effectively capture the scale-dependent joint covariance structure of multivariate wavelet coefficients as well as the associated wavelet eigenvalue structures. Extensive simulations conducted on synthetic data, modeled by operator fractional Brownian motions, the reference multivariate selfsimilarity model, permit to show that the proposed multivariate time-scale domain bootstrap based test yields the targeted significance level under the null hypothesis (all selfsimilarity exponents are equal) and to assess the power of the test for several alternative hypotheses. This analysis leads us to conclude that the proposed test for the equality of multivariate selfsimilarity exponents is effective and ready for use on (a single time series of) real data.

I. INTRODUCTION

Univariate selfsimilarity. The selfsimilarity paradigm [1] provides a relevant framework for the analysis and modeling of temporal dynamics in a wide range of applications very different in nature. Fractional Brownian motion (fBm), defined as the only Gaussian, self-similar stochastic process with stationary increments [2], [3], is probably the most well-known selfsimilarity model. It quantifies temporal dynamics in data by means of a single selfsimilarity parameter $H \in (0,1)$. The estimation of H is thus central in applications, where it is involved in various classical signal processing tasks (e.g., detection, classification, diagnosis,...), and a large body of literature has been dedicated to this topic, see, e.g., [3]–[6] and references therein. It is well documented that wavelet analysis plays a central role in this context, and has been shown to lead to theoretically well-grounded, accurate and robust estimations for H [7], [8]. Most of the work and success in applications remained limited to the modeling and analysis of univariate, scalar-valued time series so far. Yet, in an increasing number of areas, the available data naturally consist of several, jointly collected time series that call for adequate multivariate selfsimilarity models.

Selfsimilarity in multivariate time series. Operator fractional Brownian motion (ofBm) was recently proposed as a model for multivariate selfsimilarity [9]–[12]. Practically speaking, this multivariate extension of fBm can be interpreted as a mixture of M possibly correlated fBm, each with possibly distinct selfsimilarity exponents H_m, $m = 1, \ldots, M$. For this model, classical univariate-like statistical methods such as entry-wise sample covariance matrix analysis, Fourier or wavelet spectra do not yield relevant results. Recently, a statistical procedure was devised that instead relies on the scaling properties of the *eigenvalues* of the scale dependent sample wavelet covariance matrices [13]–[15]. It has been shown that this approach is effective for the joint estimation of the vector of selfsimilarity exponents $\underline{H} = (H_1, \ldots, H_M)$. The use of this novel tool in real-world applications immediately leads to a fundamental question: are observed differences in the selfsimilarity exponents $(H_1, \ldots, H_M)$ statistically significant, or are they within the natural statistical fluctuations of the estimation? A rigorous answer to this question could be achieved by formulating it as a statistical test of the hypothesis $\mathcal{H}_0 : H_1 = H_2 = \ldots = H_M$ against the alternative $\mathcal{H}_1 :$ not $\mathcal{H}_0$. Yet, so far, this has only been considered in restricted settings, non mixed ofBm in [16], mixed but non correlated ofBm in [17] or for pairs of time series only in [18], but never for ofBm with multiple, possibly many, correlated and mixed components.

Goals, contributions and outline. The present work constructs and assesses a bootstrap driven statistical test, from a single finite size observation of multivariate time series, for the equality of the selfsimilarity exponents estimated on the $M \geq 2$ components. As a first significant contribution, we expand on the wavelet estimation procedure originally proposed in [13], [15] (cf. Section II-B) to cope with the estimation bias induced by the so-named eigenvalue repulsion effect [19]–[22] (cf. Section III-A).

Building up on the wavelet domain bootstrap proposed in [18], the second main contribution consists of the bootstrap scheme, designed to preserve the time-scale joint (multivariate) dependence structure of the wavelet coefficients and thus the joint statistics of the $(H_1, \ldots, H_M)$, rather than focusing on the univariate component-wise dependence structure only as

in [18] (cf. Section III-B).

Monte Carlo simulations, conducted using OfBm synthetic data (cf. Section II-A for definitions) with $M = 4$ and 10 components, and for several sample sizes, are used to show that the proposed multivariate wavelet domain bootstrap procedure reproduces well the targeted test significance level under the null hypothesis (cf. Section IV-B), and to study the power of the proposed test under several alternative hypothesis (cf. Section IV-C). Our results demonstrate that the proposed test effectively permits to test for the equality of selfsimilarity across components from a single observation of multivariate data. OfBm synthesis, selfsimilarity exponent estimation and test procedures are implemented by the authors and will be made available at the time of publication.

II. MULTIVARIATE SELFSIMILARITY

A. Model: operator fractional Brownian motion

For simplicity, we make use of a subclass of time reversible ofBm, see [9]–[12] for an introduction of ofBm in all generality. It is a natural multivariate extension of fBm consisting of a multivariate Gaussian self-similar process with stationary increments.

Let $X \triangleq \{X_{H_1}(t), \ldots, X_{H_M}(t)\}_{t \in \mathbb{R}}$ be a collection of M fBm components, each characterized by their selfsimilarity exponents $\underline{H} = (H_1, \ldots, H_M)$, $0 < H_1 \leq \ldots \leq H_M < 1$. These components are possibly correlated according to a pointwise $M \times M$ covariance matrix Σ_X. If we denote by σ_m^2 and $\rho_{m,m'}$ the variances of each component and their respective correlation coefficients, the entries of Σ_X can be written as $(\Sigma_X)_{m,m'} = \sigma_m \sigma_{m'} \rho_{m,m'}$. It is in general not possible to choose $\underline{H}$ and Σ_X independently, see [11].

Further, let denote W a real-valued and invertible $M \times M$ matrix (the mixing matrix). Then, the M-variate stochastic process, referred to as ofBm, is defining as a linear mixture of these M correlated fBm:

$$Y \triangleq \{Y_1^{\underline{H}, \Sigma_X, W}(t), \ldots, Y_M^{\underline{H}, \Sigma_X, W}(t)\}_{t \in \mathbb{R}} \triangleq$$
$$W \{X_{H_1}(t), \ldots, X_{H_M}(t)\}_{t \in \mathbb{R}} = WX. \quad (1)$$

It can be shown that (1) implies that Y has the following key properties: i) its increments are stationary and ii) it satisfies the multivariate selfsimilarity relation:

$$\forall a > 0: \quad \{Y_1^{\underline{H}, \Sigma_X, W}(t), \ldots, Y_M^{\underline{H}, \Sigma_X, W}(t)\}_{t \in \mathbb{R}}$$
$$\overset{\text{fdd}}{\asymp} \{a^{\underline{H}}(Y_1^{\underline{H}, \Sigma_X, W}(t/a), \ldots, Y_M^{\underline{H}, \Sigma_X, W}(t/a))\}_{t \in \mathbb{R}}, \quad (2)$$

where $\overset{\text{fdd}}{\asymp}$ stands for the convergence of finite dimensional distributions, $a^{\underline{H}} \triangleq \sum_{k=0}^{+\infty} \log^k(a) \underline{H}^k / k!$, and $\underline{H} \triangleq W \text{diag}(\underline{H}) W^{-1}$ denotes the so-named selfsimilarity exponent matrix. When W is diagonal, the multivariate selfsimilarity relation simplifies to component-wise selfsimilarity relations as in [23].

B. Wavelet based joint estimation of H_m, $m = 1, \ldots, M$

The central goal of selfsimilarity analysis is to estimate the vector of exponents $\underline{H} = (H_1, \ldots, H_M)$ from the M-variate time series Y. It is only when W and thus the selfsimilarity matrix $\underline{H}$ are diagonal that each component Y_m is a fBm and $\underline{H}$ can be inferred with conventional, univariate-like approaches [16], [23]. Otherwise, each component is a mixture of fBms, and such estimations are strongly biased [13], [15]. To obtain consistent estimates for $\underline{H}$ in all generality, a statistical procedure based on the multivariate discrete wavelet transform was proposed [13], [15]. Multivariate wavelet coefficients are defined as $(D(2^j, k)) \triangleq D_Y(2^j, k) = (D_{Y_1}(2^j, k), \ldots, D_{Y_M}(2^j, k))$, $\forall k \in \mathbb{Z}$, $\forall j \in \{j_1, \ldots, j_2\}$, with $\forall m \in \{1, \ldots, M\}$: $D_{Y_m}(2^j, k) = \langle 2^{-j/2} \psi_0(2^{-j}t - k) | Y_m(t) \rangle$ and ψ_0 is a mother wavelet [24], [25]. The procedure relies on first estimating wavelet coefficient $M \times M$ sample covariance matrices at each scale 2^j independently:

$$S(2^j) \triangleq \frac{1}{n_j} \sum_{k=1}^{n_j} D(2^j, k) D(2^j, k)^*, \quad w = 1, \ldots, 2^j, \quad (3)$$

with N the sample size, j the octave and $n_j = 2^{-j} N$. Second, eigen-decompositions of these $S(2^j)$ are performed at each scale independently, yielding eigenvalues $\{\lambda_1(2^j), \ldots, \lambda_M(2^j)\}$. Then, it was shown in [13], [15] that the wavelet estimator $(\hat{H}_1^0, \ldots, \hat{H}_M^0)$ for $(H_1, \ldots, H_M)$, defined as weighted linear regressions of $\log_2 \lambda_m(2^j)$ against analysis scales $2^{j_1}, \ldots, 2^{j_2}$,

$$\hat{H}_m^0 = \left(\sum_{j=j_1}^{j_2} w_j \log_2 \lambda_m(2^j) \right) / 2 - \frac{1}{2}, \quad \forall m = 1, \ldots, M, \quad (4)$$

with w_j classical weights satisfying $\sum_j j w_j = 1$ and $\sum_j w_j = 0$ (cf. [8]), converges asymptotically in probability, $N \to \infty$, under mild assumptions, $(\hat{H}_1^0, \ldots, \hat{H}_M^0) \overset{P}{\to} (H_1, \ldots, H_M)$.

III. MULTIVARIATE WAVELET DOMAIN DE-BIASED ESTIMATION AND BOOTSTRAP HYPOTHESIS TEST

A. De-biased estimation

The estimation procedure described above shows very satisfactory practical performance but can suffer from a bias in estimation in certain situations that impairs testing the equality in H. It is indeed well known that eigenvalue decomposition applied to estimated covariance matrices is impaired by the so-called *repulsion effect*: the gap between estimated eigenvalues is on average larger than the gap between actual values, and more importantly, two equal eigenvalues are not estimated equal. This bias is all the more pronounced as the sample size available for covariance estimation is small. In wavelet transforms, the number of wavelet coefficients actually available for covariance estimation (cf. Eq. 3) decreases as scale increases. This yields a scale dependent bias in the estimation of $\lambda_m(2^j)$ that results in a bias in $\hat{H}_m^0$.

To reduce such bias, one first original contribution of the present work is to use for wavelet coefficient covariance matrix estimation the same number of wavelet coefficients at each scale. Specifically, a reference scale $2^J = 2^{j_2}$ is chosen here as the upper limit of the range where the linear fit is performed. For scale 2^j, a collection of 2^{j-j_2} covariance matrices are estimated from non-overlapping wavelet coefficients snapshots of equal size $n_{j_2} = 2^{-J}N$, for $w = 1, \ldots, 2^{j-j_2}$:

$$S^{(w)}(2^j) \triangleq \frac{1}{n_{j_2}} \sum_{k=1+(w-1)n_{j_2}}^{wn_{j_2}} D(2^j,k)D(2^j,k)^*. \quad (5)$$

The eigenvalues of each of these matrices $S^{(w)}(2^j)$, denoted as $\{\lambda_1^{(w)}(2^j), \ldots, \lambda_M^{(w)}(2^j)\}$, have similar repulsion at all scales $j \in \{j_1, \ldots, j_2\}$. Their logarithms are averaged

$$\vartheta_m(2^j) \triangleq 2^{j_2-j} \sum_{w=1}^{2^{j-j_2}} \log_2(\lambda_m^{(w)}(2^j)) \quad (6)$$

and a new wavelet estimator $(\hat{H}_1, \ldots, \hat{H}_M)$ for $(H_1, \ldots, H_M)$ is defined by means of weighted linear regressions of $\vartheta_m(2^j)$ against analysis scales $2^{j_1}, \ldots, 2^{j_2}$ as

$$\hat{H}_m = \left(\sum_{j=j_1}^{j_2} w_j \vartheta_m(2^j) \right) /2 - \frac{1}{2}, \quad \forall m = 1, \ldots, M. \quad (7)$$

B. Multivariate wavelet domain bootstrap test

Asymptotic normality and χ^2 statistics. It was shown in [13]–[15], that under mild hypothesis, the wavelet estimator $(\hat{H}_1^0, \ldots, \hat{H}_M^0)$ was characterized by jointly Gaussian multivariate statistics. Numerical simulations suggest that the new wavelet estimator proposed here, $(\hat{H}_1, \ldots, \hat{H}_M)$ for $(H_1, \ldots, H_M)$ also displays jointly Gaussian multivariate statistics. Such empirical evidence will be theoretical studied elsewhere and jointly Gaussian multivariate statistics are assumed to hold for the rest of the present work.

Given joint normality for the selfsimilarity exponents $\underline{H} = (H_1, \ldots, H_M)$, a natural statistic to test the null hypothesis $\mathcal{H}_0$ ($H_1 = H_2 = \ldots = H_M$) reads:

$$T = (\hat{\underline{H}} - \langle \hat{\underline{H}} \rangle \mathbf{1}_M)^T \Sigma_{\hat{\underline{H}}}^{-1} (\hat{\underline{H}} - \langle \hat{\underline{H}} \rangle \mathbf{1}_M) \quad (8)$$

where $\Sigma_{\hat{\underline{H}}}$ denotes the $M \times M$ covariance matrix of $\hat{\underline{H}}$, while $\langle \hat{\underline{H}} \rangle$ stands for the average across components: $\langle \hat{\underline{H}} \rangle = 1/M \sum_{m=1}^M \hat{H}_m$ and $\mathbf{1}_M = (1, \ldots, 1)$. Asymptotic joint normality of $\hat{\underline{H}} = (\hat{H}_1, \ldots, \hat{H}_M)$ implies that the test statistic T follows a χ^2 distribution with $M - 1$ degrees of freedom under the null hypothesis $\mathcal{H}_0$.

The key issue is that the matrix $\Sigma_{\hat{\underline{H}}}$ is a priori unknown and therefore needs to be replaced with an estimate (see [13]–[15] for theoretical elements). Because it is intended here that the proposed test can be put at work from a single observation of the multivariate data, $\Sigma_{\hat{\underline{H}}}$ cannot be estimated by averaging across realizations, thus calling for a bootstrap procedure [26] in the spirit of those developed in [18], [22], [27].

Wavelet domain bootstrap. To take advantage of the short term temporal and across-scale dependencies of the wavelet coefficients, the bootstrap procedure is implemented in the wavelet domain rather than in the time domain. The vector wavelet coefficients $D(2^j, k)$, $k = 1, \ldots, n_j$ are resampled jointly (i.e., in a multivariate way) by a block-bootstrap procedure which preserves their scale-time dependence structure, as opposed to a resampling for each component independently, which would not preserve the cross-component dependencies [28]. Technically, for each scale 2^j, R block bootstrap resamples $D_j^{*(r)} = (D^{*(r)}(2^j, 1), \ldots, D^{*(r)}(2^j, n_j))$, $r = 1, \ldots, R$ are drawn with replacement, from $\lceil \text{card}(Y)/L_B \rceil$ overlapping blocks of size L_B, $(D(2^j, k), \ldots, D(2^j, k + L_B - 1))$, $k = 1, \ldots, n_j$.

From each bootstrap sample of wavelet coefficients $D_j^{*(r)}$, bootstrap estimates $S^{*(r)}(2^j)$, $\vartheta_m^{*(r)}(2^j)$ and $\hat{H}_m^{*(r)}$ are produced using Eqs. (5-7). Finally, from the R vector bootstrap samples $\hat{\underline{H}}^{*(r)} = (\hat{H}_1^{*(r)}, \ldots, \hat{H}_M^{*(r)})$, $r = 1, \ldots, R$, a covariance matrix $\hat{\Sigma}_{\hat{H}^*}$ is estimated that is expected to approximate the true covariance matrix of $\underline{H} = (H_1, \ldots, H_M)$.

Bootstrap test formulation. From the estimates $\hat{\underline{H}}$ and the bootstrap covariance matrix $\hat{\Sigma}_{\hat{H}^*}$, the following test statistics is constructed:

$$T^* = (\hat{\underline{H}} - \langle \hat{\underline{H}} \rangle \mathbf{1}_M)^T \hat{\Sigma}_{\hat{H}^*}^{-1} (\hat{\underline{H}} - \langle \hat{\underline{H}} \rangle \mathbf{1}_M). \quad (9)$$

If the covariance structure is well approximated by the block-bootstrap procedure, the bootstrap statistic T^* should asymptotically be well approximated by a χ^2 distribution (as verified by numerical results in Section IV-B), hence yielding the following test decision:

$$d_\alpha = 1 \quad (\mathcal{H}_0 \text{ rejected}): \quad T^* \geq F_{\chi^2(M-1)}^{-1}(1 - \alpha), \quad (10)$$

with $F_{\chi^2(M-1)}$ the chi-square cumulative distribution function with $M - 1$ degrees of freedom and α the significance level.

IV. TEST PERFORMANCE ASSESSMENT

A. Monte Carlo Simulation

To assess the accuracy of the proposed bootstrap test procedure and actual reproduction of a targeted significance level by the test under the null hypothesis, and to quantify the power of the test under several representative alternative hypotheses, we conduct Monte Carlo simulations with synthetic data.

Specifically, the numerical simulations make use of $N_{MC} = 1000$ independent realizations of synthetic M-variate ofBm, with $M = 4$ or $M = 10$ components, and with three different sample sizes $N = 2^{14}, 2^{16}$, and 2^{18}. The covariance matrix Σ_X is set to a Toeplitz matrix, with off-diagonal entries all set to $\rho = 0.5$. The mixing matrix W is chosen randomly in the space of invertible $M \times M$ matrices. Both matrices are kept fixed for all the experiments.

Under the null hypothesis $\mathcal{H}_0$, the selfsimilarity exponents are set to $H_1 = \ldots = H_M = 0.7$. Several significance levels $\alpha \in \{0.005, 0.01, 0.05, 0.1, 0.15\}$ are studied.

For the alternative hypothesis, two scenarios are tested. In Scenario1, there are only two different values of H, denoted H_1 and H_2, amongst the M components. In Scenario1a,
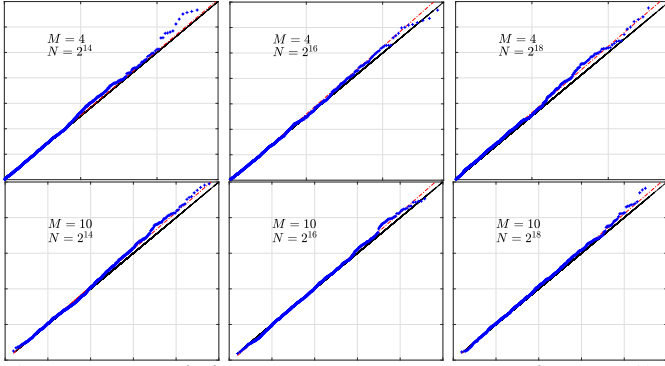
Fig. 1: **Test statistics under the null hypothesis.** Quantile-quantile plot of the bootstrap statistic T^* against a theoretical χ^2 distribution with $M-1$ degrees of freedom.

Fig. 2: **Distributions of p-values under the null hypothesis.** Empirical distributions of the p-values of the proposed bootstrap test compared to the Uniform Distribution.

Fig. 3: **Significance levels.** Empirical significance levels (obtained as averages of the bootstrap test decisions against realizations) against targeted ones.

the two groups of components corresponding with H_1 and H_2 have equal size $M_1 = M_2 = M/2$; in Scenario1b, they have unbalanced sizes $(M_1, M_2) = (1, 3)$ for $M = 4$, $(M_1, M_2) = (2, 8)$ for $M = 10$. Further, $H_1 = 0.5$ is fixed and H_2 is varied in the interval $[0.5, 0.9]$. In Scenario2, there are four groups of components of sizes $(M_1, M_2, M_3, M_4) = (2, 3, 2, 3)$ ($M = 10$) and $(M_1, M_2, M_3, M_4) = (1, 1, 1, 1)$ ($M = 4$) with selfsimilarity exponents ($H_1 = 0.5, H_2 = H_1 + \Delta H, H_3 = H_1 + 2\Delta H, H_4 = H_1 + 3\Delta H$), $\Delta H \in [0, 0.13]$. The significance level is set to $\alpha = 0.05$.

For estimation, the multivariate wavelet transform is computed with the least asymmetric Daubechies3 wavelets, estimation scales are set to $j_1 = 6$ and $j_2 = \log_2(N) - 5$. For the test, $R = 500$ bootstrap resamples are used with $L_B = 6$ corresponding to the size of the Daubechies3 mother-wavelet time support.

B. Test behavior under the null hypothesis

Fig. 1 reports quantile-quantile plots of the empirical distribution of T^* obtained from the N_{MC} independent realizations against the theoretical chi-square distribution with $M - 1$ degrees of freedom. It shows, for three different samples sizes N and two different number of components M, an excellent agreement, hence validating the heuristic assumption in Section III-B that the test statistic T^* follows χ^2 distribution.

Furthermore, Fig. 2 shows that the empirical distributions of the corresponding p-values are approximately uniform, as should theoretically be the case under the null hypothesis $\mathcal{H}_0$.

To finish with, Fig. 3 shows, for the three different samples sizes N, and the two different number of components M, that the achieved significance levels, obtained as averages of the binary decisions of the bootstrap test across realizations, reproduce very satisfactorily the targeted test significance.

Overall, these results strongly suggest that, for large ranges of sample sizes N and numbers of components M, the bootstrapped statistics T^* reproduce well the distribution of the true statistics T and follow well a χ^2 distribution with $M - 1$ degrees of freedom, and that the proposed bootstrap procedure test is effective in controlling the significance level (type-I error) of the test.
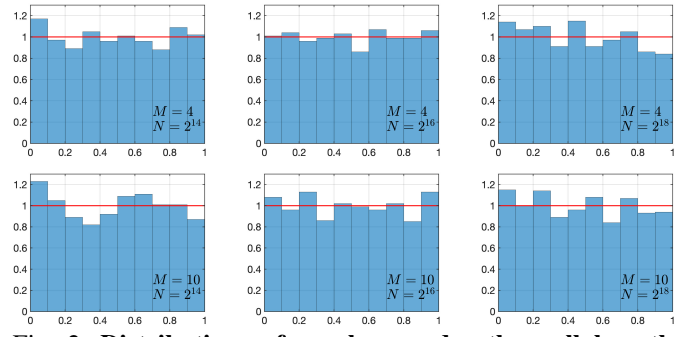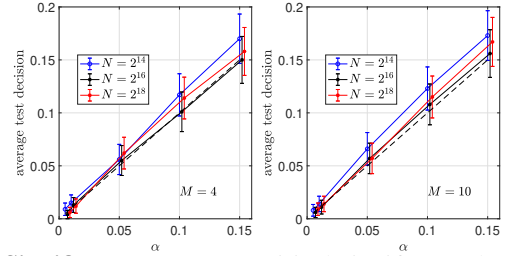
C. Power of the test under different alternative hypothesis

Fig. 4 report the empirical assessment (by average across independent realizations of the test decisions) of the power of the proposed bootstrap-based χ^2 tests as functions of $\Delta H = |H_2 - H_1|$, for Scenario1 and Scenario2, respectively. In both scenarios, and as expected, the power of the test increases with sample size and with ΔH. From an elementary study of the χ^2 statistics T, and given that the variance in the estimation of H essentially decreases as $1/N$ (see [13]–[15]), it can be deduced that the power of the test increases as $\Delta H \cdot \sqrt{N}$.

In addition, Fig. 4 (top plots) shows that, for a scenario with two different H and a fixed ΔH, the test power decreases when the number of components M increases.

Further, and interestingly, comparing the results of Scenario1a and 1b shows that, for a fixed ΔH, test powers are larger when the groups of components with same H have equal sizes, compared to when groups of components have unbalanced sizes, which can be explained as follows. Under alternative hypothesis, the test statistics T^* follows a noncentral χ^2 distribution with $M-1$ degrees of freedom, with a noncentrality parameter that explicitly depends on the values of the true values $H_1, \ldots, H_M$, as well as on the covariance and mixing matrices Σ_X and W. The test power, directly stemming from the distance between the central and noncentral χ^2 distributions, both with $M - 1$ degrees of freedom, is thus fully determined by this noncentrality parameter. For instance, for Scenario1, calculations not detailed here show that the noncentrality parameter is essentially proportional to
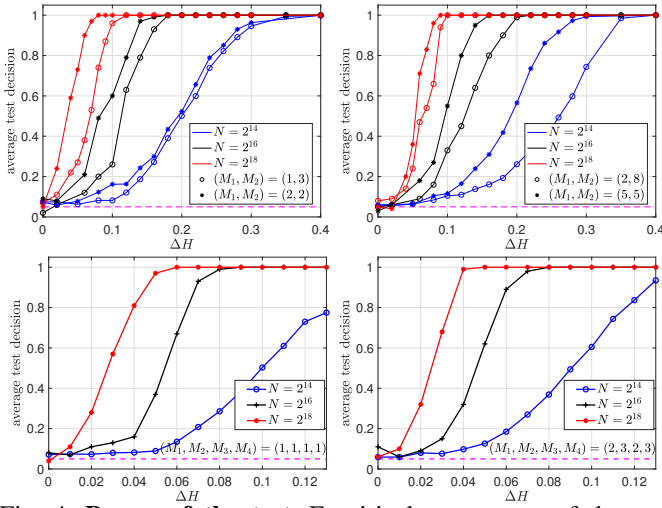
Fig. 4: **Power of the test.** Empirical assessment of the test powers, obtained as average of the decision of the bootstrap test, as a function of ΔH, for $M = 4$ (left) and $M = 10$ (right), for different sample sizes, and with equal or different sizes in the groups of equal components, for Scenario1 (top) and Scenario2 (bottom). Significance level is set to $\alpha = 0.05$ (magenta dashed line).

$(M_1 \times M_2)/(M_1 + M_2)$ and thus decreases when $|M_2 - M_1|$ increases, and so does the test power.

Fig. 4 (bottom plots) finally shows that the test also has significant power when there are more than two different values of H.

V. Conclusions and perspectives

We have devised and assessed a multivariate wavelet domain bootstrap procedure that permits to test, from a finite size single observation of a multivariate time series, whether or not the selfsimilarity exponents are equal. Monte Carlo simulations, based on operator fractional Brownian motion, the reference process to model multivariate selfsimilarity, enabled us to show that the procedure is effective in controlling an a priori prescribed significance level and that the proposed test benefits from significant power under several alternative hypothesis. The proposed test is ready for application to real-world multivariate time series. OfBm synthesis, selfsimilarity exponent estimation and testing procedures are implemented by the authors and will be made available at the time of publication. The work opens interesting perspectives related to i) the assessment of how many different values for H actually exist amongst the M components and ii) the study of large dimensional cases, where the number of components M and sample size N become comparable in magnitude.

Acknowledgment

References

[1] G. Samorodnitsky and M. Taqqu, *Stable non-Gaussian random processes*. New York: Chapman and Hall, 1994.

[2] B. B. Mandelbrot and J. W. van Ness, "Fractional Brownian motion, fractional noises and applications," *SIAM Reviews*, vol. 10, pp. 422–437, 1968.

[3] V. Pipiras and M. S. Taqqu, *Long-Range Dependence and Self-Similarity*. Cambridge University Press, 2017, vol. 45.

[4] J. Beran, *Statistics for Long-Memory Processes*. Chapman & Hall/CRC, Boca Raton, 1994.

[5] J. Beran, Y. Feng, S. Ghosh, and R. Kulik, *Long-Memory Processes: Probabilistic Properties and Statistical Methods*. Springer-Verlag, New York, 2013.

[6] H. Wendt, P. Abry, and S. Jaffard, "Bootstrap for empirical multifractal analysis," *IEEE Signal Processing Mag.*, vol. 24, no. 4, pp. 38–48, 2007.

[7] P. Flandrin, "Wavelet analysis and synthesis of fractional Brownian motion," *IEEE Trans. Info. Theory*, vol. 38, no. 2, pp. 910–917, 1992.

[8] D. Veitch and P. Abry, "A wavelet-based joint estimator of the parameters of long-range dependence," *IEEE Trans. Info. Theory*, vol. 45, no. 3, pp. 878–897, 1999.

[9] M. Maejima and J. D. Mason, "Operator-self-similar stable processes," *Stochastic Processes and their Applications*, vol. 54, no. 1, pp. 139–163, 1994.

[10] J. D. Mason and Y. Xiao, "Sample path properties of operator-self-similar Gaussian random fields," *Theory of Probability & Its Applications*, vol. 46, no. 1, pp. 58–78, 2002.

[11] G. Didier and V. Pipiras, "Integral representations and properties of operator fractional Brownian motions," *Bernoulli*, vol. 17, no. 1, pp. 1–33, 2011.

[12] ——, "Exponents, symmetry groups and classification of operator fractional Brownian motions," *Journal of Theoretical Probability*, vol. 25, pp. 353–395, 2012.

[13] P. Abry and G. Didier, "Wavelet estimation for operator fractional Brownian motion," *Bernoulli*, vol. 24, no. 2, pp. 895–928, 2018.

[14] G. Didier, H. Helgason, and P. Abry, "Demixing multivariate-operator self-similar processes," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proces. (ICASSP)*, Brisbane, Australia, 2015, pp. 3671–3675.

[15] P. Abry and G. Didier, "Wavelet eigenvalue regression for n-variate operator fractional Brownian motion," *Journal of Multivariate Analysis*, vol. 168, pp. 75–104, November 2018.

[16] H. Wendt, G. Didier, S. Combrexelle, and P. Abry, "Multivariate Hadamard self-similarity: testing fractal connectivity," *Physica D*, vol. 356-357, pp. 1–36, 2017.

[17] P. Abry, G. Didier, and H. Li, "Two-step wavelet-based estimation for Gaussian mixed fractional processes," To appear in *Statistical Inference for Stochastic Processes*, pp. 1–61, 2018.

[18] H. Wendt, P. Abry, and G. Didier, "Wavelet domain bootstrap for testing the equality of bivariate self-similarity exponents," in *Proc. IEEE Workshop Statistical Signal Proces. (SSP)*, Freiburg, Germany, June 2018.

[19] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An Introduction to Random Matrices*, ser. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2010, vol. 118.

[20] T. Tao, *Topics in Random Matrix Theory*. American Mathematical Society, 2012, vol. 132.

[21] J. Yao, S. Zheng, and Z. Bai, *Sample Covariance Matrices and High-Dimensional Data Analysis*. Cambridge University Press, 2015.

[22] H. Wendt, P. Abry, and G. Didier, "Bootstrap-based bias reduction for the estimation of the self-similarity exponents of multivariate time series," in *IEEE Int. Conf. Acoust., Speech, and Signal Proces. (ICASSP)*, Brighton, UK, May 2019.

[23] P.-O. Amblard and J.-F. Coeurjolly, "Identification of the multivariate fractional Brownian motion," *IEEE Trans. Signal Proces.*, vol. 59, no. 11, pp. 5152–5168, 2011.

[24] I. Daubechies, *Ten Lectures on Wavelets*. SIAM, 1992, vol. 61.

[25] S. Mallat, *A Wavelet Tour of Signal Processing*. San Diego, CA: Academic Press, 1998.

[26] A. M. Zoubir and D. R. Iskander, *Bootstrap Techniques for Signal Processing*. Cambridge University Press, 2004.

[27] H. Wendt, P. Abry, and S. Jaffard, "Bootstrap for empirical multifractal analysis," *IEEE Signal Proces. Mag.*, vol. 24, no. 4, pp. 38–48, 2007.

[28] S. N. Lahiri, *Resampling Methods for Dependent Data*. New York: Springer, 2003.